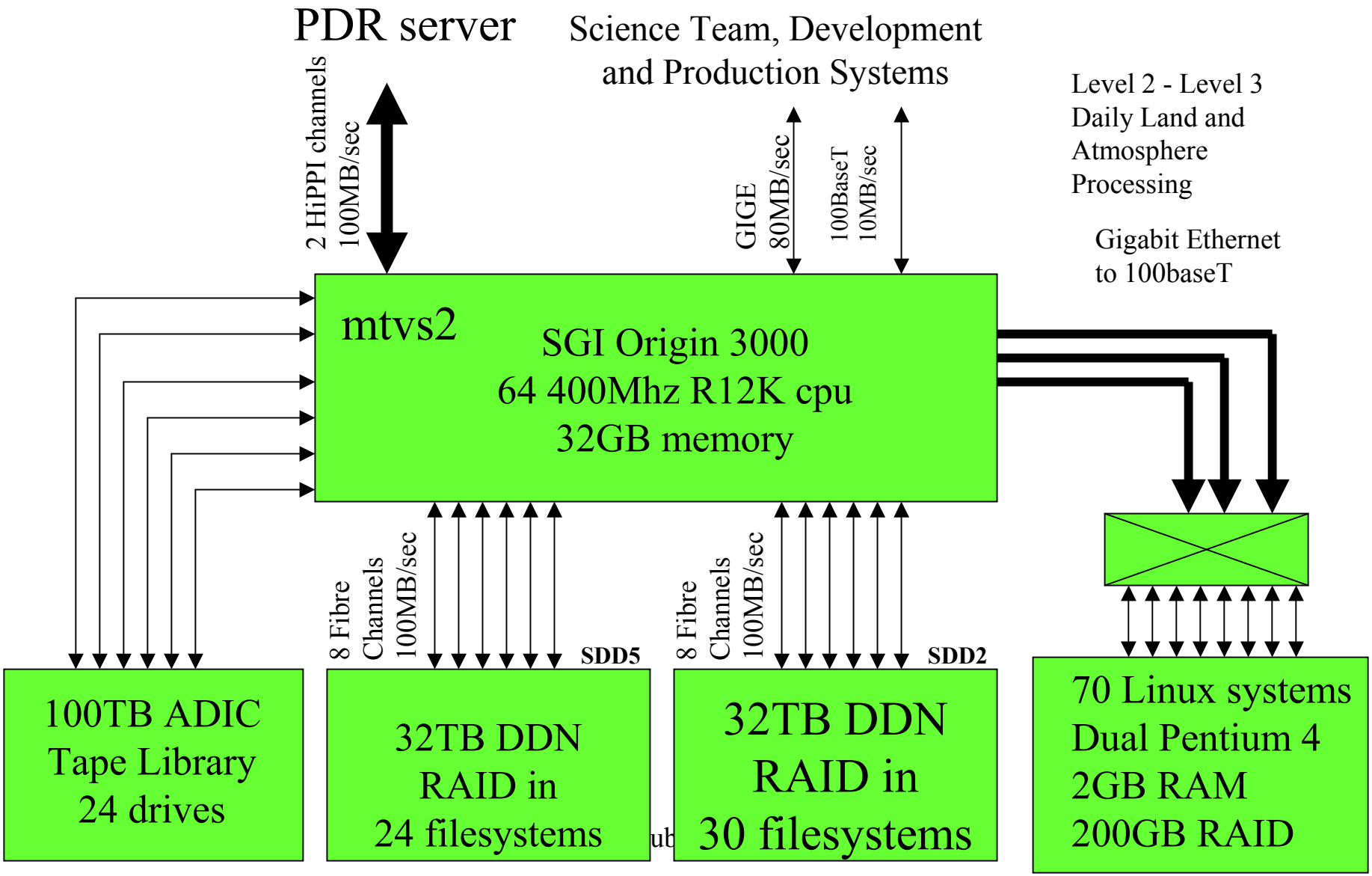
# Reuse lessons learned

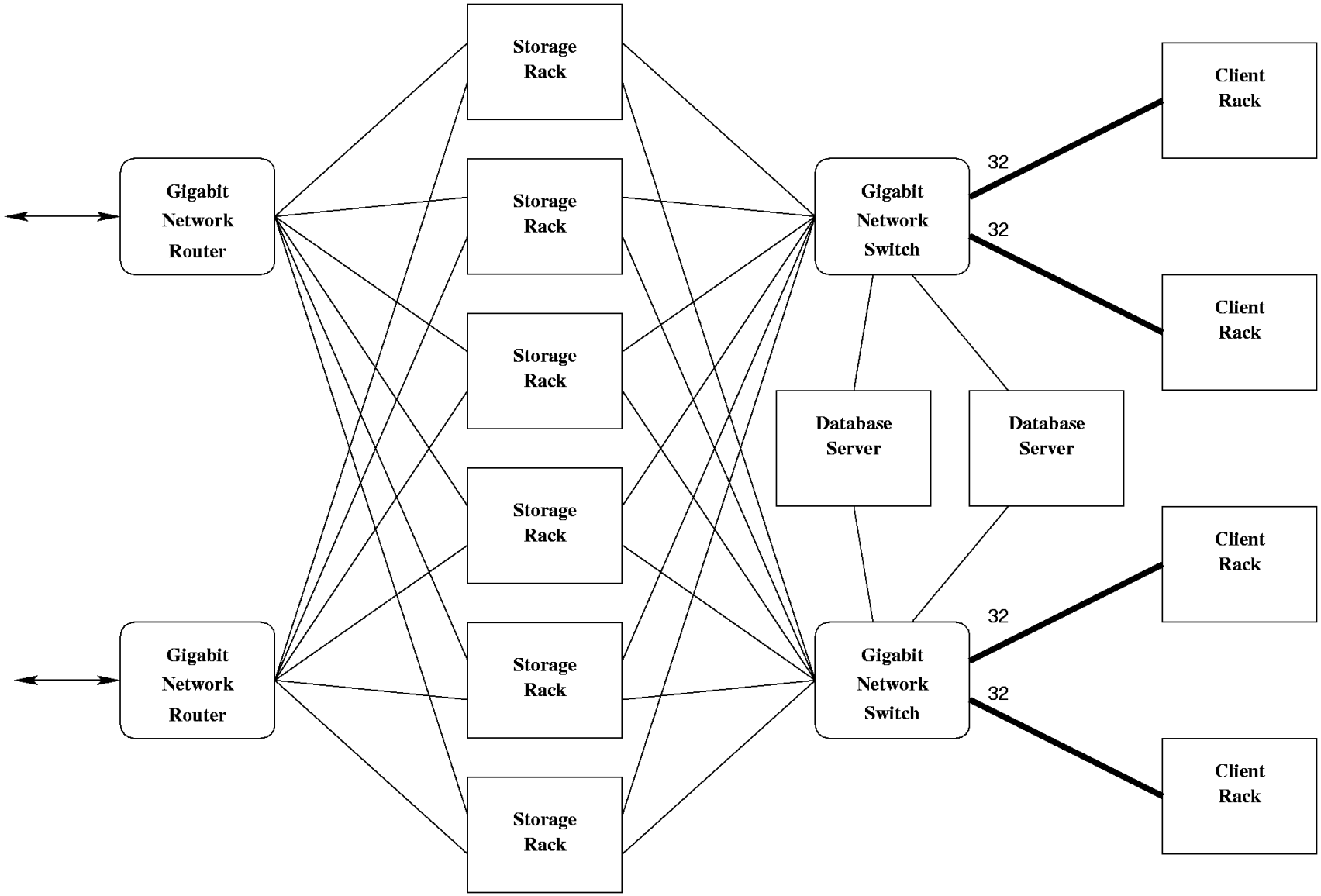
Ed Masuoka
MODIS Science Data Support Team

# What is MODAPS

- MODIS SIPS
  - Close integration of development, operations, science
- Can produce all MODIS products
  - Produces and distributes L2 and L3 products to DAACs
  - Making >2.25TB/day during current reprocessing
- Supports search/order and distribution to MODIS Science Team (>300GB/day)
- Runs on a variety of systems
  - Laptops, SGI Origin family, Commodity PCs (Linux)

# mtvs2 reprocessing string




Storage Rack
Storage Rack
Storage Rack
Storage Rack
Storage Rack
Storage Rack
Gigabit Network Router
Gigabit Network Router
Gigabit Network Switch
Gigabit Network Switch
Database Server
Database Server
32
32
32
32
Client Rack
Client Rack
Client Rack
Client Rack

# MODAPS History

- SeaWiFS reused for MODIS Emergency Backup System 1997-1998 (.2X)
- 1999 V1 MODAPS (1X) still used SeaWiFS as a foundation
- MODAPS rewritten to handle MODIS production at higher rates (>6X) with support for commodity hardware running Linux
- OSIPS (production system for OMI) based on MODAPS with PostgresSQL replacing Sybase and with the PGEs framework made more general

# Other data system reuse in EOS

- "DAACWare" (V0 DAAC Scheduling and Data Management Software) reused for ICESAT SIPS and MLS SIPS by Mark Sherman's (Raytheon) team
- TSDIS for TRMM will evolve to produce data for GPM under Eric Stocker's team

# What do we gained from reuse

- Ability to rapidly field data systems
- Stability of underlying data system
- One or more of the items below:
  - Cost savings (~60%)
  - Time to develop or incorporate the “next great things” needed by stakeholders
  - More resources for science team and Q/A

# Reuse of elements from EOSDIS

- B0 Data Model
  - Defines metadata about products
- Science Data Processing Toolkit (SDPTK)
  - Isolate science software from operating system
- HDF-EOS
  - Self describing format for data products

# Reuse of Open Source Software

- Linux - Operating System
- Perl - Production Rule Scripts
- Openssl - Encrypted network transactions
- Apache - Web server
- Postfix - Mail
- PostgresSQL - Relational DBMS
  - MODAPS uses Sybase, OSIPS uses PostgresSQL
- CVM - Configuration Management
  - MODAPS uses ClearCase, OSIPS uses CVM

# Why was software reused?

- EODIS elements - implement common standards and were mandated by ESDIS Project
- Open source software - free and performs a set of needed functions well
- Data systems - reused by the same developers for a series of missions.

# Factors influencing reuse

- Schedule and near term/long term costs
- Familiarity with the  data system
- Pride of authorship/ownership
- Fit to the application
  - features, scalability, adaptability ...
- Modularity
- Standards/Tools that promote interoperability
- Support for development team

# Suggestions

- Maintain a common data model
  - Start with EOSDIS data model - improve Q/A info
  - Make it easy to extend
  - Develop primer for new teams
- Maintain a common self-describing product format
  - HDF-EOS works for us
  - metadata in products are invaluable for future teams
- Enable easy addition or replacement of services
  - improved search/order, enhanced distribution…

# Suggestions

- Maintain information about reusable components
  - Software, documentation, support
  - Virtual test-bed for interoperable components (in facilities from existing missions)
- Consider guidelines to enable software portability to allow for improvements in hardware
  - Porting PGEs to Linux gave MODAPS significant gains in performance for relative small costs
    - Made possible by SDP-Toolkit and rules to insure portability
- Explore reuse in areas like quality assurance
  - Both knowledge capture and specific tools